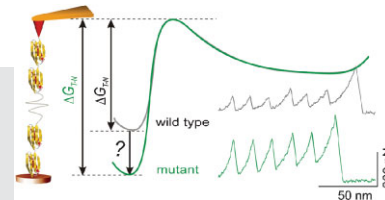


DOI: 10.1002/adfm.200800480

‘Mechanical Engineering’ of Elastomeric Proteins: Toward Designing New Protein Building Blocks for Biomaterials**

By Hongbin Li*

Elastomeric proteins are subject to stretching force under biological settings and play important roles in regulating the mechanical properties of a wide range of biological machinery. Elastomeric proteins also underlie the superb mechanical properties of many protein-based biomaterials. The developments of single molecule force spectroscopy have enabled the direct characterization of the mechanical properties of elastomeric proteins at the single molecule level and led to the new burgeoning field of research: single protein mechanics and engineering. Combined single molecule atomic force microscopy and protein engineering efforts are well under way to understand molecular determinants for the mechanical stability of elastomeric proteins and to develop methodologies to tune the mechanical properties of proteins in a rational and systematic fashion, which will lead to the ‘mechanical engineering’ of elastomeric proteins. Here the current status of these experimental efforts is discussed and the successes and challenges in constructing novel proteins with tailored nanomechanical proteins are highlighted. The prospect of employing such engineered artificial elastomeric proteins as building blocks for the construction of biomaterials for applications ranging from material sciences to biomedical engineering is also discussed.



1. Introduction

Elastomeric proteins are an important class of mechanical proteins that are subject to stretching force under physiological conditions.^[1–8] They function as molecular springs to provide tissues with extensibility, elasticity, and mechanical strength. Some elastomeric proteins are also biomaterials of superb mechanical properties.^[2] For example, spider dragline silk is the best known fibrous material that outperforms any manmade high performance fibrous material.^[9,10] Resilin, an elastomeric protein found in specialized regions of the cuticle of most insects, has amazing resilience properties and plays important roles in insect flight.^[11] The mechanical properties of these amazing elastomeric proteins are ‘encoded’ in their three-dimensional structures as well as their unique organization into materials. At the level of materials science, extensive experimental efforts have been well under way for decades to understand the design and organization of these materials, and mimic the design of such biological materials to engineer novel biomaterials with a wide variety of mechanical properties.^[2,11] Parallel to these efforts on the

macroscopic level, investigations into the molecular design principles of these elastomeric proteins are also under way, thanks to the recently developed single molecule force spectroscopy based techniques.^[4,7,12,13] Using single molecule force spectroscopy techniques, it has become possible to stretch individual elastomeric proteins and study their mechanical properties and structure–function relationship one molecule at a time. Understanding the molecular details of the design of elastomeric proteins is not only important for elucidating the biophysical principles that underlie a wide variety of biological processes,^[14,15] but also may illustrate new design principles for biomaterials and pave the way to design novel elastomeric proteins with well-defined mechanical properties using bottom-up approaches.^[16] These efforts will also help to use these novel elastomeric proteins for nanobiotechnological applications.

Depending upon their biological functions, elastomeric proteins can be made of largely unstructured proteins,^[3,17] such as elastin, to provide high entropic elasticity and extensibility, or tandem modular proteins that consist of individually folded protein domains, which convey high toughness and serve as shock-absorbers, or both.^[5] The latter is found in a wide variety of biological systems, ranging from muscle fiber^[5] to biological adhesive employed by abalones.^[6] For example, the giant muscle protein titin is an elastomeric protein made of hundreds of individually folded Ig domains and largely unstructured unique sequences.^[5] The overall mechanical properties of such elastomeric proteins are largely determined by the mechanical properties of the constituting protein domains. Extensive single molecule atomic force microscopy (AFM) studies and molecular dynamics simulations have been carried out to determine the mechanical properties of a wide range of elastomeric proteins and to illustrate their underlying molecular design principles.^[18–32] It

[*] Prof. H. Li
Department of Chemistry
The University of British Columbia
2036 Main Mall
Vancouver, BC, V6T 1Z1 (Canada)
E-mail: hongbin@chem.ubc.ca

[**] The author is grateful for the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chairs Program, the Canada Foundation for Innovation (CFI), the Canadian Institutes of Health Research (CIHR), Michael Smith Foundation for Health Research and the University of British Columbia. Mr. Yi Cao and Qing Peng are also thanked for their critical reading of the manuscript and their help in preparing Figure 1.

was discovered that stretching force can trigger sequential mechanical unfolding of individually folded domains in tandem modular proteins. Such 'modular' unfolding provides a unique mechanism to dissipate energy and conveys high toughness to elastomeric proteins, making them perfect shock-absorbers.^[4,7,33] The elastic properties of individual protein domains are combined collectively to determine the overall mechanical properties of elastomeric proteins. For example, single molecule AFM studies on titin have provided insights into the molecular mechanism of how the passive elasticity of muscle is finely regulated by the collective mechanical properties of the constituting folded Ig-like domains as well as random coil-like sequences of titin.^[33–36] Understanding the molecular determinants of the mechanical stability of proteins will not only illustrate some fundamental principles governing the mechanical stability of proteins, but also provide the possibility to engineer novel elastomeric proteins with tailored nanomechanical properties that will serve as building blocks for the bottom up construction of novel biomaterials with applications in material sciences and biomedical engineering. Inspired by naturally occurring elastomeric proteins, researchers have started to explore and develop new methodologies to tailor the mechanical properties of proteins in a rational way with the aim to exploit the engineered artificial elastomeric proteins for specific nanomechanical applications.^[37–43] Over the last ten years, significant progress has been made in this new burgeoning area of research. In this Feature Article, an overview of the recent advances in the field of single protein mechanics is provided, with a particular focus on the studies of engineering elastomeric proteins with tailored nanomechanical properties. Although many studies in this area are closely associated with the investigation of mechanical unfolding dynamics of proteins, the focus here is only on the discussion of the mechanical property aspect of these studies.

2. Investigating the Mechanical Properties of Proteins Using Single Molecule AFM

2.1. Principles and Operation Modes of Single Molecule AFM

Historically, the mechanical properties of proteins were generally deduced from macroscopic measurements of the

mechanical properties of protein-based materials. It was not until the late 1990s that the mechanical properties of individual proteins could be directly measured at the single molecule level.^[4,12,13] The development of single molecule force spectroscopy techniques made it possible to mechanically manipulate individual polymer chains and measure their force–extension relationships at the single molecule level and with pico-Newton resolution in force and nanometer precision in extension. Among the single molecule force spectroscopy techniques, single molecule AFM is especially suitable to measure the mechanical properties of proteins because of its superb spatial resolution ($\text{\AA}$), force sensitivity (~ 10 pN), force range (from ~ 10 pN to $\sim$ nN), and no need of specific chemical immobilization in most of the studies.^[18] The combination of protein engineering techniques with single molecule AFM^[44] makes it possible to investigate the mechanical design of elastomeric proteins in great detail, which gives rise to the burgeoning field of research: single protein mechanics. For single molecule AFM studies, a polyprotein made of identical tandem repeats of the protein of interest is usually constructed using protein engineering techniques.^[18,44,45] Elastic properties of the constructed polyprotein are then measured using AFM in one of the two operation modes: force-extension mode^[4,44,46] and force-ramp mode.^[47–49] Figure 1 shows schematics of single molecule AFM experiments in which an engineered polyprotein is stretched and unravelled by AFM in these two modes.

2.2. Mechanical Stability of Proteins

The mechanical stability of a given protein can be defined as the force needed to unravel a protein at a given pulling velocity or loading rate. The mechanical unfolding force of a given protein depends on two intrinsic parameters that characterize the mechanical unfolding energy diagram of proteins: the free energy barrier, $\Delta G_{\text{T-N}}$, for mechanical unfolding and the distance Δx_u between the folded state and the mechanical unfolding transition state.^[50–52] Figure 2 shows a schematic free energy diagram for the mechanical unfolding of a protein in the absence of stretching force. Upon being subject to a stretching force, the unfolding free energy diagram tilts, which results in a decrease of the unfolding free energy barrier and



Hongbin Li is an Assistant Professor of Chemistry and a Canada Research Chair (Tier II) in molecular nanoscience and protein engineering in the Department of Chemistry, University of British Columbia, Canada. He obtained his Bachelor degree in Polymer Engineering from Tianjin University in 1993, and his Ph.D. in polymer chemistry and physics from Jilin University, P.R. China, in 1998. From 1996 to 1997, he was a joint-training Ph.D. student in Prof. Hermann E. Gaub's group in Ludwig-Maximilians-Universität München, Germany. He worked with Prof. Julio M. Fernandez as a postdoctoral research fellow at the Mayo Foundation (1999–2002) and then as an Associate Research Scientist at Columbia University (2002–2004). In 2004, he joined the faculty in the Department of Chemistry at University of British Columbia. He is a recipient of a Michael Smith Foundation for Health Research Career Investigator Award and was a Peter Wall Institute for Advanced Studies Early Career Scholar. His research interests include elastomeric proteins, single molecule atomic force microscopy, protein folding/unfolding dynamics, protein engineering, and biomaterials.

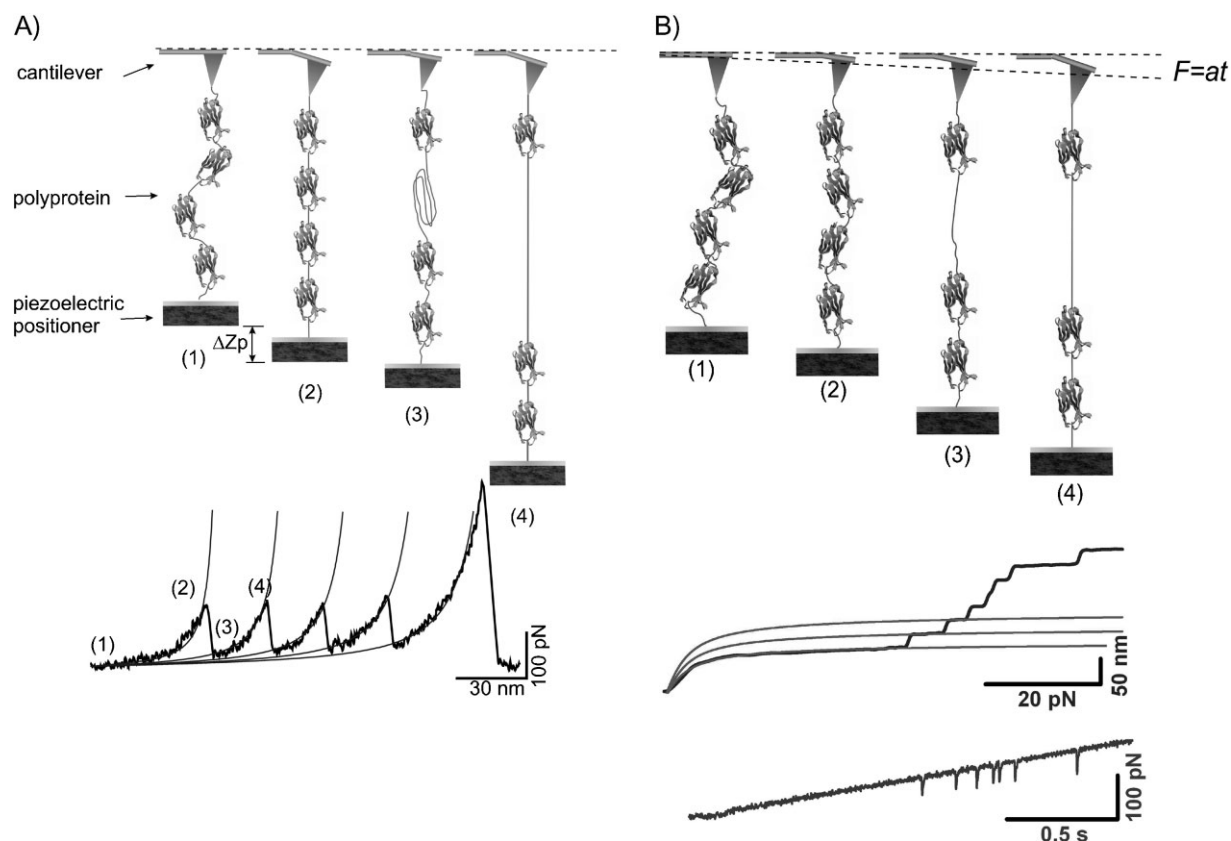


Figure 1. Using single molecule atomic force microscopy to probe the mechanical properties of single proteins. In a typical single molecule AFM experiment, a tandem modular protein, which is deposited onto a glass cover slip, is picked up by the AFM tip and stretched between the AFM tip and the solid substrate, which is mounted onto a high precision piezoelectric positioner. There are two different operation modes of AFM: force-extension mode (A) and force-ramp mode (B). A) A schematic of a force-extension measurement on a polyprotein. In force-extension mode of AFM, the two ends of the protein are stretched apart at a constant velocity by moving the piezoelectric positioner away. The force can be measured from the deflection of the AFM cantilever. Stretching a tandem modular protein in the force-extension mode results in force-extension curves of the characteristic saw-tooth pattern appearance of force peaks. The individual sawtooth peak corresponds to the sequential unravelling of individual domains in the tandem modular protein. The unfolding force is a measure of the mechanical stability of the protein domains. As the piezoelectric positioner moves away to increase the end-to-end distance of the molecule (from state 1 to state 2), the protein generates a restoring force following the worm-like-chain model of polymer elasticity. Upon domain unfolding, the contour length of the protein increases and the force acting on the cantilever is relaxed. Further extension again results in the increase of force (state 4). The last peak in the force-extension curve represents the extension of the fully unfolded tandem modular protein prior to its detachment from the AFM tip or substrate. B) A schematic of a force-ramp experiment on a polyprotein. In this mode, the stretching force F increases linearly as a function of time ($F = at$, where a is the ramp rate and t is time), and the end-to-end distance of a single tandem modular protein is measured as a function of F . The mechanical unfolding of individual domains gives rise to the staircase appearance of the resultant extension-force curves. The extension-force curve (middle panel) is characterized by the step-wise elongation of the end-to-end distance and can be well described by the worm-like-chain model of polymer elasticity (grey lines). The measured force signal as a function of time is shown as the bottom panel. Due to the limited frequency response, transient relaxation of the force correlates with the domain unfolding event and is shown as spikes. Adapted from [49]. Copyright 2006 Elsevier.

the increase of the unfolding rate constant.^[50–52] The dependence of the unfolding rate constant $\alpha(F)$ on the applied stretching force can be described by the Bell–Evans model:^[50]

$$\alpha(F) = A \cdot e^{-\frac{\Delta G_{T-N} - F \cdot \Delta x_u}{k_B T}} = \alpha_0 \cdot e^{\frac{F \cdot \Delta x_u}{k_B T}} \quad (1)$$

where A is the pre-exponential factor, k_B is Boltzmann constant, and T is temperature in Kelvin. When the protein is subject to a force ramp, which changes linearly following the relationship of $F = a \cdot t$ (Fig. 1B), where a is the loading rate, and t is time, it can be calculated that the most probable unfolding force F_u follows the following

equation:^[52]

$$F_u = \frac{k_B T}{\Delta x_u} \cdot \ln \left(\frac{a \cdot \Delta x_u}{\alpha_0 \cdot k_B T} \right) \quad (2)$$

It is evident that ΔG_{T-N} (or α_0) and Δx_u collectively influence the mechanical stability of proteins. It can be shown that a smaller spontaneous unfolding rate constant α_0 (e.g., larger unfolding free energy barrier) and smaller unfolding distance Δx_u favor a higher unfolding force. Thus, upon mutation, the change in the unfolding force of the protein can not be

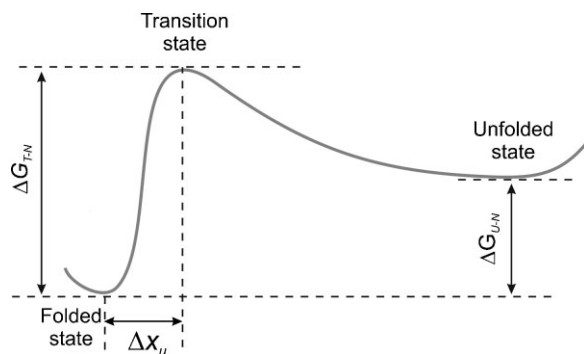


Figure 2. Schematic free energy diagram for the mechanical unfolding of proteins. ΔG_{T-N} denotes the free energy barrier for the mechanical unfolding reaction, and Δx_u denotes the unfolding distance between the native state and the mechanical unfolding transition state. ΔG_{U-N} denotes the thermodynamic stability of proteins, i.e., the free energy difference between the unfolded and folded states of proteins.

predicted from ΔG_{T-N} (or α_0) or Δx_u alone. For example, an increased mechanical unfolding energy barrier ΔG_{T-N} upon mutation does not necessarily lead to a higher mechanical stability as the information about Δx_u is also needed. In contrast, the loading rate $\dot{\alpha}$ influences the mechanical unfolding force of proteins in a straightforward way: the higher the loading rate, the higher the unfolding force.

In comparison with force-ramp experiments, most single molecule AFM experiments are carried out in the force-extension mode, e.g., the end-to-end distance of the protein is increased linearly with time by moving the piezoelectric positioner away at a constant velocity (Fig. 1A).^[4] The unfolding force of proteins in force-extension measurements is dependent upon pulling speed: the faster the pulling speed, the higher the unfolding force. It is worth noting that, although Equation (1) is universal in describing the mechanical unfolding of proteins, the mechanical unfolding force measured in force-extension experiments cannot be predicted by a simple analytical formula, which is in contrast with force-ramp experiments.

2.3. Mechanical Stability versus Thermodynamic Stability and Kinetic Stability

It is clear that, mechanical stability, e.g., mechanical unfolding force, is directly correlated with the kinetic stability ΔG_{T-N} , which is the free energy barrier for the mechanical unfolding reaction along the mechanical unfolding pathway defined by the stretching force.^[35] In contrast, the thermodynamic stability is the free energy difference between the unfolded and native states of proteins (ΔG_{U-N}). Therefore, the mechanical stability does not correlate with thermodynamic stability, which was experimentally verified by single molecule AFM experiments.^[35] Furthermore, although mechanical stability is correlated with mechanical kinetic stability, mechanical stability is generally not correlated with the chemical kinetic stability of proteins.^[53–56] This is because

the mechanical unfolding and chemical unfolding may follow different pathways and do not necessarily coincide with each other.^[53,55,57] The coincidence between mechanical and chemical unfolding rate constants were observed in a few proteins, such as Ig domains from titin^[33,35,44,58] and GB1 domain from protein G,^[56] however, the coincidence disappeared when point mutation was introduced.^[53–55] It is evident that, although mechanical unfolding and chemical unfolding rate constants may be similar for some proteins, such a coincidence can not and should not be generalized to other proteins of interests, even to their point mutants. The general lack of correlation between mechanical stability and chemical kinetic stability/thermodynamic stability makes it impossible to predict the mechanical stability based on available kinetic and thermodynamic data on proteins *a priori*.

2.4. Anisotropic Nature of Mechanical Resistance

Another distinct feature of mechanical stability, or mechanical resistance, is anisotropy. Thermodynamic stability and chemical kinetic stability are global properties of a protein, and each protein has a well-defined thermodynamic and kinetic stability. In contrast, mechanical unfolding proceeds along a predefined reaction coordinate determined by the vector of the stretching force. Depending upon the direction along which the protein is mechanically unravelled, the mechanical stability of the same protein can be different. Single molecule AFM experiments and molecular dynamics simulations have provided direct supporting evidence.^[45,59–61] Therefore, the mechanical response of a protein to a stretching force is anisotropic. Here the discussion is focussed on the mechanical stability of proteins being stretched from their N- and C-termini only.

3. Expanding the Toolbox of Elastomeric Proteins to Include Non-mechanical Proteins

Naturally occurring elastomeric proteins are subject to a stretching force under their biological settings. To perform their biological functions under such stressful biological environments, significant mechanical stability is likely to be a prerequisite for such naturally occurring elastomeric proteins. Extensive single molecule force spectroscopy studies have revealed that indeed these naturally occurring elastomeric proteins are generally mechanically stable and can resist a stretching force up to a few hundred piconewtons under laboratory experimental conditions. The progress in understanding naturally occurring elastomeric proteins has been reviewed in numerous articles,^[14,15,18–20,22,62] and interested readers are referred to these references. Among the naturally occurring elastomeric proteins, the giant muscle protein titin is the most extensively studied model protein.^[14,33] Although naturally occurring elastomeric proteins constitute a rich toolbox for nanomechanical engineering, they may not be

sufficient for engineering multi-functional nanomechanical devices and materials.^[63] Thus it becomes necessary to explore the potential utility of proteins other than naturally occurring elastomeric proteins in nanomechanical engineering.

In contrast to the naturally occurring elastomeric proteins, a large number of proteins are not subject to a stretching force under their biological settings, and a stretching force is unlikely to be an evolutionary pressure for them. We refer to these proteins as non-mechanical proteins. Questions naturally arise from such distinctions: Is mechanical stability a property unique to naturally occurring elastomeric proteins? Can non-mechanical proteins be mechanically stable and used for nanomechanical purposes? To answer these questions, extensive single molecule AFM studies have been carried out on a variety of non-mechanical proteins in the hope to significantly expand the toolbox of elastomeric proteins.^[30,37,38,40,42,60,64–68] An early single molecule AFM study on a non-mechanical protein barnase^[65] suggested that proteins that are not selected for mechanical functions may not resist force in the same way as mechanical proteins. Considering the non-mechanical nature of barnase, this result was not surprising. Searching non-mechanical proteins that are of significant mechanical stability has continued, and much expanded experimental efforts have revealed a much brighter prospect for using non-mechanical proteins for mechanical applications. A series of

mechanically stable non-mechanical proteins have been identified and characterized since.^[38–40,42,43] Inspecting the structures of naturally occurring elastomeric proteins revealed that mechanical proteins do not contain unique covalent structures that provide mechanical strength, instead, mechanical proteins use the same set of non-covalent interactions that determine the overall three-dimensional structures and thermodynamic stability of proteins to determine the mechanical stability of proteins. Consequently, the mechanical stability of a protein must depend on its optimal use and arrangement of known interactions, such as hydrogen bonds, electrostatic interactions, hydrophobic interactions, etc. Therefore, non-mechanical proteins could exhibit significant mechanical stability if their three-dimensional structures adopt arrangements of key non-covalent interactions that are similar to those of natural mechanical proteins.

Steered molecular dynamics (SMD) simulations of the mechanical unfolding of natural mechanical proteins revealed that protein topology plays an important role in determining the mechanical stability of proteins,^[24,25,69–72] and single molecule AFM experiments led to a similar conclusion.^[15,18] Figure 3 (middle panel) shows some of the well-characterized naturally occurring elastomeric proteins that are of significant mechanical stability. In these proteins, a common feature is that the two terminal force-bearing β -strands are arranged in

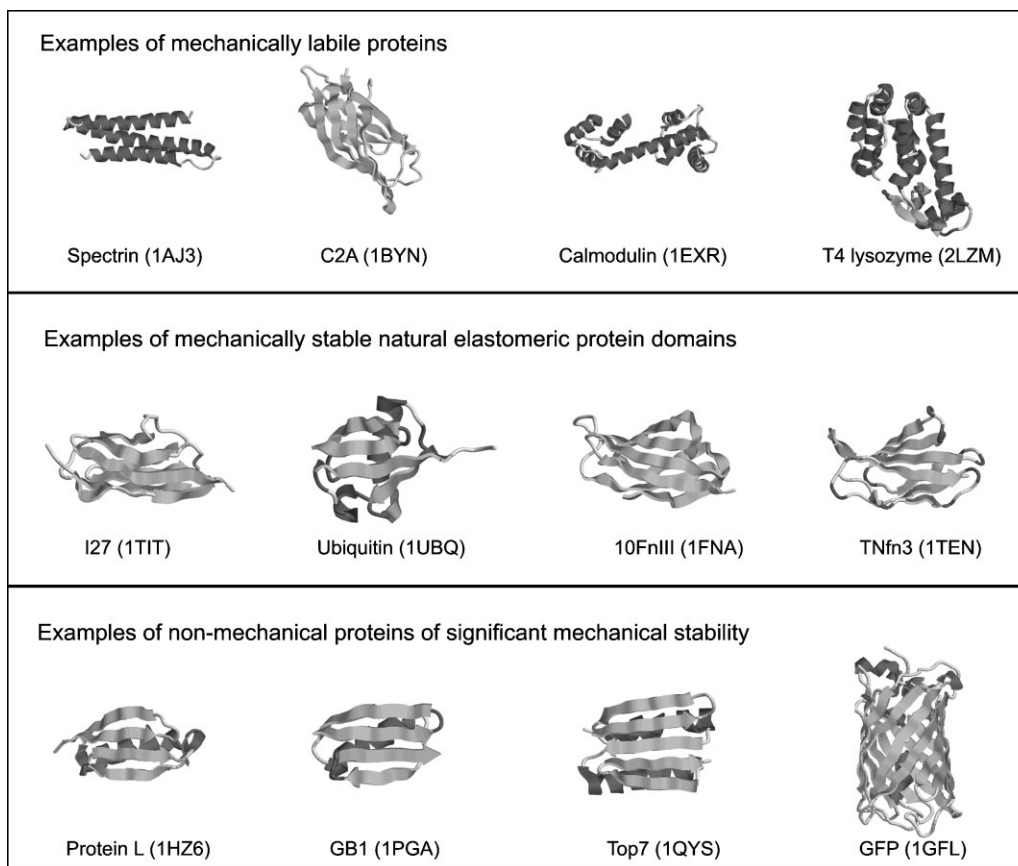


Figure 3. Three-dimensional structures of representative proteins that have been investigated using single molecule AFM.

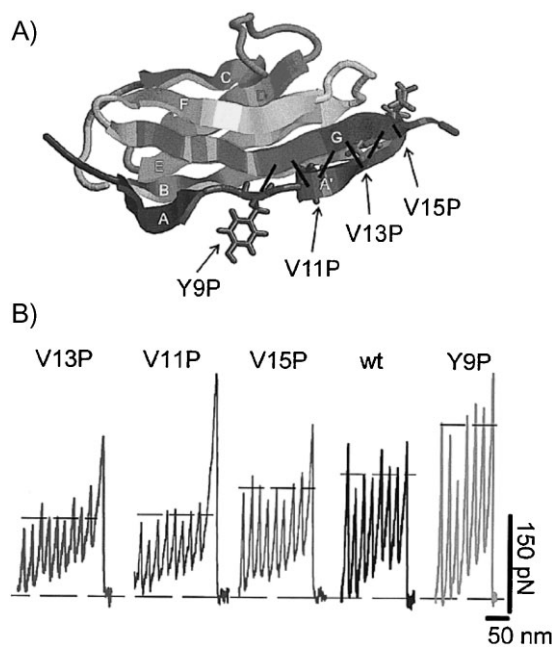


Figure 4. Point mutations in the mechano-active site of I27 alter its mechanical stability. A) Cartoon diagram showing the β -sandwich structure of the I27 module and the amino acids that were substituted by proline residues. Black bars indicate the six backbone hydrogen bonds linking the A' and G β -strands that are predicted to be the mechano-active site and hold the key to the mechanical stability of I27. B) The force extension relationships for the wt I27 and I27 proline mutant polypeptides. The mutations V11P, V13P, and V15P decrease the force required to unfold the I27 module. By contrast, the mutation Y9P increases this force. Reproduced with permission from [53]. Copyright 2000 Macmillan Publishers Ltd.

parallel and the N- and C-termini are pointing in opposite directions. Such an arrangement of the force-bearing β -strands constitutes a shear topology. The arrangement of A' and G β -strands in I27 is a typical example of the shear topology (Fig. 4A). The result of such a shear topology is that the interactions, such as backbone hydrogen bonds and hydrophobic interactions, which hold the two force-bearing strands together have to be unravelled more or less concurrently in order to extend the two termini of the protein. Therefore, these interactions serve as a mechanical clamp to resist mechanical unfolding and form the molecular basis for the mechanical stability of the protein.^[69,70] In contrast, mechanically labile proteins (Fig. 3, upper panel) do not possess such a shear topology. In these mechanically labile proteins, similar interactions that connect the two force-bearing motifs, being β -strands or α -helices, can be unravelled sequentially,^[18] which leads to low mechanical stability. If this view is correct, non-mechanical proteins in principle can display or be engineered to display significant mechanical stability just like their mechanical counter parts, provided that these non-mechanical proteins possess, either by nature or by engineering, desired shear topology of their force-bearing β -strands.

Experimental efforts have validated such reasoning, so do some molecular dynamics simulations.^[72] Using shear topology as a search criterion, a series of non-mechanical proteins of

significant mechanical stability have been successfully identified and characterized. Protein L,^[38] GB1 domain of protein G,^[42,43] and Top7,^[40] are three representative examples of the identified non-mechanical proteins. Similar to ubiquitin,^[59] GB1 and protein L belong to the β -grasp fold, with the shear topology arrangement of the terminal force-bearing β -strands. However, GB1 and protein L do not have known mechanical functions under their biological settings. Single molecule AFM studies showed that protein L^[38] and protein G^[42,43] exhibit a significant mechanical stability and unfold at forces of ~ 130 and ~ 180 pN, respectively, which are comparable to that of elastomeric proteins, such as the I27 domain from titin^[44] and ubiquitin.^[59]

Shear topology is an important structural feature for mechanically stable proteins. Is the direct contact between the two force-bearing β -strands in shear topology a necessary condition? Is there any other protein fold that is potentially mechanically stable? With these questions in mind, we investigated the mechanical unfolding of a small protein Top7. Top7 is a *de novo* designed novel protein using *ab initio* methods with a novel protein fold that has not been observed in any naturally occurring proteins.^[73] The computationally designed structure of Top7 matches the experimentally determined high-resolution structure with atomic level accuracy, thus Top7 serves as a perfect example of a non-mechanical protein. In Top7, the two force-bearing β -strands are not in direct contact with each other, but spaced by a third β -strand (bottom panel of Fig. 3 and Fig. 5A). Despite such structural variations, Top7 was shown to be mechanically stable^[74] with an average unfolding force of ~ 150 pN. These results demonstrated that direct contact of force-bearing strands is not a necessary condition for mechanical stability in a shear topology. Furthermore, the Top7 fold is distinct from the Ig-like fold and β -grasp fold, thus represents a novel mechanically stable protein fold.^[74] Although the shear topology seems to be a general feature for significant mechanical stability, it does not exclude the possibility that proteins without shear topology are also mechanically stable. Recent studies on green fluorescent protein (GFP) unveiled such complexity.^[39,45,75] GFP is a β -barrel protein. Upon stretching from different directions by engineered cysteine residues, no shear topology is apparent in many of the pulling directions and yet GFP exhibits significant mechanical stability.^[45] Moreover, it is likely that there are different levels of molecular determinants for the mechanical stability of proteins, and shear topology is just one of the many. Other unknown factors might well bring new insight/surprise to our understanding. The recent discovery that ankyrin is mechanically stable and shows unusual elastic behavior is a perfect example in this regard.^[23,27]

Since non-mechanical proteins can have unique properties in both their functions and dynamics, the inclusion of non-mechanical proteins into the toolbox of elastomeric proteins will not only increase the number elastomeric proteins that mimic the mechanical properties of those natural ones, but can also bring some unique properties that are not possessed by

natural elastomeric proteins. For example, the inclusion of GFP can introduce optical features in an elastomeric protein which may find unique applications in constructing optical-mechano sensors. In addition, artificial elastomeric proteins can also entail mechanical properties that may outperform those of natural ones. For example, our recent experiments showed that an artificial polypeptide made of GB1 folds much faster and can recover its mechanical stability more efficiently than any elastomeric protein that has been studied to date.^[43] In addition, this artificial polypeptide also exhibits very little mechanical fatigue after long periods of continuous stretching–relaxation cycles.

Moreover, the finding that Top7 exhibits significant mechanical stability is of particular significance. Top7 was *de novo* designed by Baker and co-workers in 2003 as an effort toward designing proteins with arbitrarily chosen three-dimensional structures. Top7 was shown to have a novel sequence and have a protein fold that has not been sampled by nature.^[73] The finding that Top7 is mechanically stable^[40] unveils the feasibility to computationally *de novo* design proteins of novel topology to possess tailored nanomechanical properties, although the design of Top7 was not intended for mechanical purposes as such. The next challenges for the field of protein mechanics will be to computationally design novel proteins of well-defined mechanical stability and such efforts will test our current understanding of molecular determinants of the mechanical stability for proteins.

In parallel to these single molecule AFM experiments, recently an ambitious simulation effort^[72] was undertaken to investigate the mechanical stability of all the proteins that have known three dimensional structures in the protein structural database Protein Data Bank. Using Go-like models, Cieplak and co-workers simulated the mechanical unfolding of the proteins in the Protein Data Bank and measured their mechanical stability by computer simulation. A large number of proteins were predicted to be mechanically stable, with a big fraction of these potential candidates being non-mechanical proteins. An interesting observation that emerged from such predictions is that most of the mechanically stable proteins do share the shear topology of force-bearing β -strands. These studies demonstrate the great potential of non-mechanical proteins to achieve desirable mechanical properties, and will greatly expand the toolbox of mechanically stable proteins for nanomechanical applications.

4. Rational Tuning of the Mechanical Stability of Proteins

Tuning the mechanical stability of proteins rationally is not only important to understand the molecular determinants of mechanical stability, but also key to use designed elastomeric proteins for material science and biomedical applications. Mechanical stability is determined by the unfolding distance Δx_u as well as the free energy difference ΔG_{T-N} between the mechanical unfolding transition state and native state. To

tune the mechanical stability of proteins, it is necessary to change the relative energetics of the native state and mechanical unfolding transition state. Despite the distinct difference between mechanical stability and thermodynamic stability, tuning the mechanical stability of proteins is analogous in many ways to tuning (enhancing) the thermodynamic stability of proteins in the field of enzyme engineering. Enhancing the thermodynamic stability of proteins can be achieved in two different directions: rational design (both computationally and experimentally) and laboratory based directed evolution. Similarly, tuning the mechanical stability of proteins can be also classified into these two categories. In this part, the discussion will focus on tuning the mechanical stability of proteins by design. Experimental efforts along this direction have been well under way towards developing rational and systematic methodologies to tune the mechanical stability of proteins. Given that molecular determinants of the mechanical stability remain largely unestablished, most of the efforts are still trial-and-error based. In a few special cases, rational tuning of the mechanical stability has become possible. Here this progress is discussed along two independent directions: 1) tuning the mechanical stability by chemical modification of proteins (Sections 4.1 to 4.4), and 2) tuning the mechanical stability of proteins by physical (or environmental) means (Sections 4.5 to 4.6).

4.1. Tuning the Mechanical Stability by Modifying the Mechano-active Site

As demonstrated by single molecule AFM and SMD simulations, local topology and interactions are critical to the mechanical stability of proteins as they can form a mechanical clamp to provide the necessary mechanical resistance to unfolding. For example, the AB and A'G regions of I27 are shown to be key to the mechanical stability of I27 (Fig. 4).^[53,55,57,69,70] The backbone hydrogen bonds that connect the A' and G strands form the mechanical clamp to resist mechanical unfolding (Fig. 4). Such critical region(s) of a mechanical protein can be considered as the mechano-active site, analogous to the active site for an enzyme. Therefore, modifying the mechano-active site of proteins using site-directed mutagenesis becomes the most natural approach to tune the mechanical stability of proteins. The first experimental studies using this approach were carried out on I27, the paradigm for single protein mechanics.^[53,57] Using site-directed mutagenesis, residues in the A' strand that are involved in the formation of key backbone hydrogen bonds in the mechano-active site A'G region were mutated to proline.^[53] It is known that proline substitution blocks the formation of backbone hydrogen bonds, introduces a bulge in the β strand, and also affects the hydrophobic packing, and hence leads to disruption of the local β -sheet structure. As expected, disruption of the mechano-active site by proline substitutions at positions Val11, Val13, and Val15 causes a significant phenotypic effect of the mechanical unfolding of I27: the unfolding forces of these I27 proline mutants were

reduced significantly and the unfolding pathway of I27 was also changed as evidenced by the large increase in the unfolding distance Δx_u between the native state and transition state. Moreover, a surprising finding was that a supposedly disruptive proline mutation at position Tyr9 increased the mechanical stability from 200 to 250 pN. The molecular origin for such a surprising mechanical stabilization effect remains mysterious, because of the lack of detailed structural information of Y9P I27. Since then, extensive studies have been carried out to investigate the phenotypical effects of point mutations on the mechanical stability of proteins, and tuning the key interactions in the mechano active site has become a widely used approach.^[34,76–79]

Despite extensive efforts to tune the mechanical stability of proteins by modifying the mechano-active site, molecular determinants of the mechanical stability of proteins remain not fully understood. As such, most of the experimental efforts lead to decreased mechanical stability of proteins, and increasing the mechanical stability of a given protein by mutating the mechano-active site proves challenging. Enhancing the mechanical stability of a given protein by point mutation remains largely trail-and-error based and success in such efforts remains rare. Y9P I27 remains to be the only point mutant that 'accidentally' did the trick!^[53] Searching for such 'accidental' successful cases will surely help lead to rational approaches to tune the mechanical stability of proteins. Moreover, it has been recognized that the coupling of the mechano-active site with the rest of the protein structure may also be a factor one needs to take into account when attempting to enhance the mechanical stability by mutating the mechano-active site. For example, recent studies showed that, despite the local attributes of mechano-active sites, mutations outside the mechano-active site may also affect the mechanical stability.^[54,74,80,81] Understanding such coupling may reveal new avenues towards rational tuning of the mechanical stability of proteins.

4.2. Rational Tuning of Mechanical Stability by Controlling the Mechanical Unfolding Pathway

Mechanical stability is a property unique to its associated mechanical unfolding pathway. Therefore, it is possible to control the mechanical unfolding pathway of a protein to achieve predefined mechanical stability. Recently, based on a model system Top7, we developed a new strategy to realize this idea to tune the mechanical stability of Top7 in a rational fashion.^[40]

Combining single molecule AFM and SMD, we discovered that Top7 unfolds by a substructure-sliding mechanism.^[40] The three-dimensional structure of Top7 is 'symmetric' with respect to the center β -strand (Fig. 5A). Hence, Top7 could unfold by two potential unfolding pathways (Fig. 5A): one being the sliding of substructure A against B/C, and the other one being the sliding of substructure C against A/B. SMD simulations revealed that the unfolding of Top7 always proceeds by the pathway of sliding substructure A

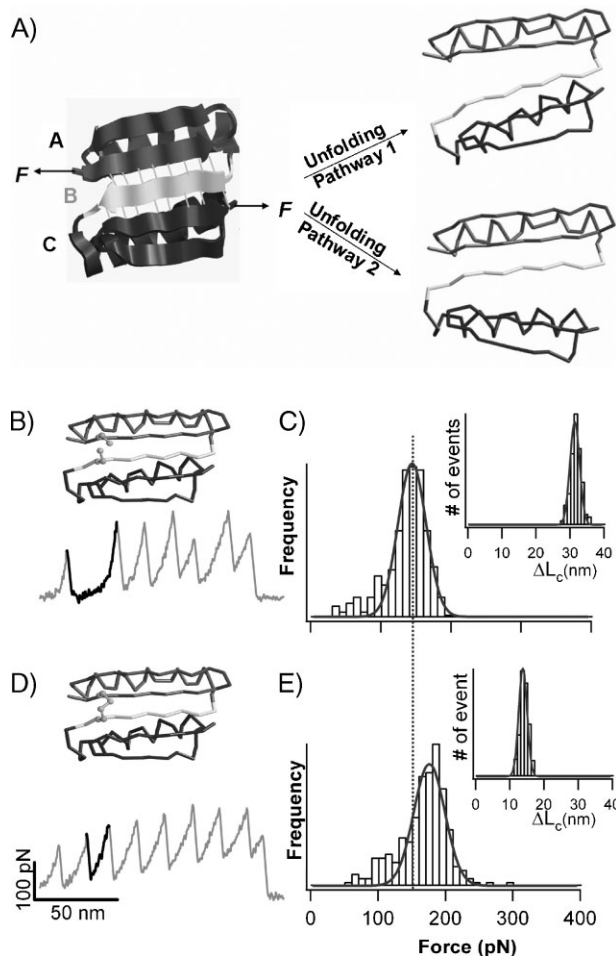


Figure 5. Tuning the mechanical stability of Top7 by redesigning its mechanical unfolding pathway. A) There are two potential unfolding pathways for Top7: the first one corresponds to the sliding of substructure A against B/C, while the second one corresponds to the sliding of substructure C against A/B. SMD simulations show that the first unfolding pathway dominates the unfolding of Top7. B–E) The formation of a disulfide bond modulates the mechanical unfolding pathway of Top7 and its mechanical stability. Force-extension curves and cartoon representations of designed Top7 mutants are shown in (B) and (D). B) Mechanical properties of reduced Q3C/T51C-Top7. In the presence of reducing agent dithiothreitol (DTT), the disulfide bond does not form. The force-extension curves show unfolding events of reduced Q3C/T51C with ΔL_c of ~ 30 nm (in black). C) The average unfolding force of reduced Q3C/T51C is 140 pN and ΔL_c is 31.0 ± 2.0 nm (Inset). D) The mechanical stability of oxidized Q3C/T51C increased due to the shifting of the unfolding pathway. Upon oxidation, Cys3 and Cys51 form a disulfide bond that covalently links strands 1 and 3, blocking the unfolding pathway of sliding substructure A against B/C. The unfolding of oxidized Q3C-T51C results in unfolding events with ΔL_c of ~ 13 nm (in black). E) The average unfolding force of oxidized Q3C/T51C is 172 pN, a ~ 30 pN increase as compared with the reduced Q3C/T51C, and ΔL_c is 13.5 ± 1.7 nm (Inset). Solid lines in C) and E) are Gaussian fits. Adapted with permission from [40]. Copyright 2007 The National Academy of Sciences of the USA.

against B/C, which suggests that the two potential unfolding pathways are not equivalent in free energy barrier and the observed pathway of sliding A against B/C is the one with a lower free energy barrier. Based on this insight from SMD simulations, it was reasoned that, if we could block the sliding

of substructure A against B/C, it would then be possible to force Top7 to unfold by the pathway of sliding A/B against C, which is the pathway of higher energy barrier. Following this reasoning, we computationally designed disulfide mutants of Top7 to covalently link substructure A with B and specifically block the unfolding pathway of a lower free energy barrier. Upon stretching the oxidized form of the Top7 disulfide mutant, Top7 unfolded by the pathway of the higher energy barrier and exhibited an increased mechanical stability. Lowering the mechanical stability of a given protein is well within the reach of current knowledge; however, it remains challenging to rationally increase the mechanical stability of a protein. The successful example of Top7 demonstrates that regulating the mechanical unfolding pathway of proteins is a unique approach towards the challenge of rational enhancing the mechanical stability of proteins. Moreover, this method can be further explored and extended to other protein systems that potentially have more than one unfolding pathway. For such purposes, FnIII domains from both fibronectin and tenascin might be excellent model systems. For example, SMD simulations predicted that the unfolding of the tenth FnIII domain from fibronectin could unfold by two distinct unfolding pathways involving unfolding intermediate states,^[82,83] which were verified later by single molecule AFM studies.^[78] One unfolding pathway is by unravelling from the N-terminus, and the other one is by unravelling from the C-terminus. It can be envisioned that a similar approach may be used to block one possible unfolding pathway of FnIII domain and force it to unfold by a different one. Thus the mechanical stability of FnIII domains can be readily tuned by controlling the choice of the mechanical unfolding pathways. We are currently testing such possibilities.

4.3. Regulating the Mechanical Stability of Proteins by Configurational Entropy

Configurational entropy plays an important role in defining the thermodynamic stability as well as the folding/unfolding kinetics of proteins. However, its role in regulating the mechanical stability of proteins remains largely unexplored. Recently, we combined single molecule AFM and protein engineering techniques to investigate the role of configurational entropy in regulating the mechanical unfolding kinetics and mechanical stability of protein GB1.^[84] We systematically elongated the length of the second loop of GB1 that connects the α -helix and the third β -strand (loop 2) by two, five, twenty four, and forty six flexible residues and measured their mechanical stability using single molecule AFM. We found that loop elongation significantly decreased the mechanical stability of GB1 and accelerated its mechanical unfolding kinetics, which is suggestive of the importance of loop2 in the mechanical unfolding pathway of GB1. Such effects can be satisfactorily explained by the loss of configurational entropy upon closing an unstructured flexible loop, and provide unique possibilities of modulating the mechanical stability of proteins by controlling the configurational entropy. Since the effective

length of the flexible loop can be modulated by disulfide bridges,^[85] we are currently exploring the use of disulfide crosslinking to achieve reversible regulation of the mechanical stability of proteins.

4.4. Tuning the Mechanical Stability by Optimizing the Packing of Hydrophobic Core

Although mechano-active sites in proteins play dominant roles in determining the mechanical stability of proteins, the overall three-dimensional structure of a given protein as well as the interactions adjacent to the mechano-active site also play important roles. As such, optimizing hydrophobic interactions, e.g., optimizing the packing of a hydrophobic core, may also lend effective ways to tune the mechanical stability of proteins. Our recent case study on a GB1 mutant, Gc3b4, shows some promise. Gc3b4 is a computationally designed mutant by Mayo and co-workers to improve its thermodynamic stability through optimizing hydrophobic core packing.^[86] Our single molecule molecular AFM studies showed that the mechanical stability of Gc3b4 increases to ~ 210 pN from ~ 180 pN for wt GB1,^[81] which suggests that it is possible to achieve improved mechanical stability by optimizing core packing. However, the molecular mechanism to achieve the improved mechanical stability remains unknown. Along the same line, a recent study on FnIII domains from tenascin and fibronectin showed that replacing the hydrophobic core of the tenth FnIII domain from fibronectin by that of the third FnIII domain from tenascin-C led to the improved mechanical stability of the hybrid FnIII domain, which highlights the potential importance of hydrophobic core packing.^[80] These studies are just a beginning, and expanded efforts along this direction will contribute to the delineation of the molecular determinants of the mechanical stability for proteins.

4.5. Enhancing the Mechanical Stability of Proteins by Ligand Binding and Protein-Protein Interactions

Ligand binding, including protein-protein interactions, is ubiquitous in nature. It is well known that ligand binding can increase the thermodynamic stability of proteins by affecting the equilibrium between the folded and unfolded states of proteins.^[87,88] Therefore, ligand binding and protein-protein interactions have been extensively exploited for protein stabilization, both in nature and in the laboratory. For example, protein complexes are perfect examples nature uses to build cellular machinery with considerable thermodynamic stability. However, mechanical unfolding pathways are not necessarily the same as the chemical/thermo unfolding pathways, and mechanical stability does not generally correlate with thermodynamic stability of protein. Hence, it was not clear whether ligand binding and protein-protein interactions could enhance the mechanical stability of proteins as they do the thermodynamic stability.

Using dihydrofolate reductase (DHFR) from Chinese hamster as a model system, Fernandez and co-workers elegantly

demonstrated that ligand binding can enhance the mechanical stability of DHFR.^[89] Using single molecule AFM and fingerprint techniques, they showed that in the absence of ligands, DHFR unfolds at very low forces (average at 27 pN) and the stretching of DHFR does not result in any characteristic mechanical unfolding signature. Instead, a featureless mechanical response that is typical of random coil-like unfolded polypeptides was often observed. Upon adding its ligand nicotinamide adenine dinucleotide phosphate (NADPH), dihydrofolate (DHF), or inhibitor methotrexate (MTX), the mechanical stability of DHFR was significantly enhanced, which resulted in unfolding events of DHFR at ~80 pN. In addition, they also found that the binding of multiple ligands to DHFR simultaneously did not result in additive stabilizing effects. These findings clearly demonstrate the feasibility of using ligand binding to enhance a protein's mechanical stability. It is interesting to note that there may exist potential complexity of using a ligand to enhance the mechanical stability of DHFR from different organisms. Two other studies showed that DHFR from mouse and *E. coli* have very different response to ligand binding.^[66,67] In these two cases single molecule AFM experiments showed that the mechanical stability of these two forms of DHFR do not show any change upon binding a ligand (MTX and NADPH).

Recently, using a small protein GB1 as a model system, we have combined protein engineering and single molecule AFM techniques to systematically investigate the effect of protein-protein interactions on a protein's mechanical stability and develop methodologies of using protein-protein interactions to enhance the mechanical stability of proteins.^[81,90] GB1 is well known for its high affinity binding to IgG antibodies. There are two GB1-binding epitopes in IgG, one is in the Fab region and the other one is in the Fc region. It is worth noting that the Fc binding site of GB1 is distant from the mechano-active site of GB1, which are the force-bearing β -strands 1 and 4. Using single molecule AFM, we demonstrated that the binding of Fc and Fab to GB1 can significantly increase the mechanical stability of GB1. As shown in Figure 6A, GB1 unfolds at ~180 pN at a pulling speed of 400 nm s⁻¹.

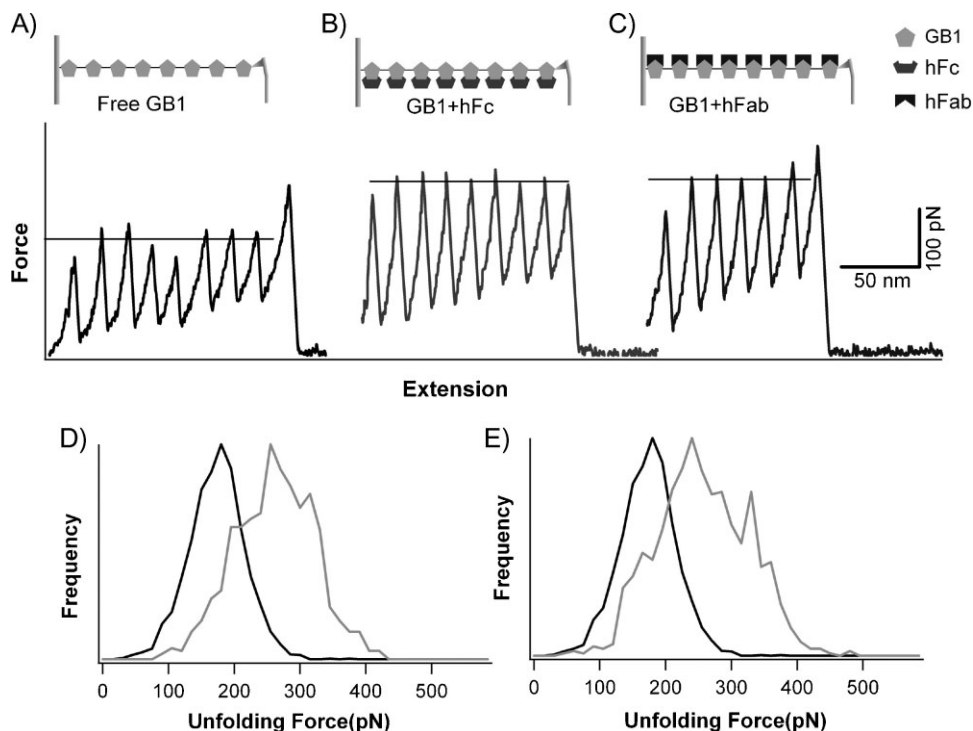


Figure 6. The mechanical stability of protein GB1 is significantly enhanced by the binding of human IgG fragments hFc and hFab. A–C) Force-extension curves of GB1 polypeptide, GB1/hFc complex, and GB1/hFab complex. Top panels show the schematics of stretching polypeptides of GB1, GB1 in complex with Fc and GB1 in complex with Fab between an AFM tip and glass substrate, respectively. Stretching polypeptide (GB1)₈ results in force-extension curves of a typical saw-tooth pattern that are characterized by unfolding forces of ~180 pN. Each individual force peak corresponds to the mechanical unfolding of an individual GB1 domain in the polypeptide. The mechanical stability of GB1 is enhanced by the binding of hFc (B) and hFab (C). When pre-equilibrated with $\sim(11\text{--}25) \times 10^{-6}$ M of hFc, the majority of GB1 domains unfold at much higher forces of ~260 pN, as indicated by the solid line, which is ~80 pN higher than that for GB1 in the absence of hFc. Similarly, in the presence of $(50\text{--}70) \times 10^{-6}$ M of hFab, GB1 unfolds at ~260 pN. D–E) Unfolding force histograms of GB1 (black), GB1/hFc complex (D, grey), and GB1/hFab complex (E, grey). It is evident that the binding of hFc and hFab to GB1 significantly enhances the mechanical stability of GB1. Adapted with permission from [81]. Copyright 2008 Elsevier.

Upon binding to Fc or Fab (Fig. 6B,C), the mechanical unfolding force of GB1 increases dramatically to ~260 pN, which indicates that the binding of Fc or Fab to GB1 has significant mechanical stabilization effects on GB1. Since the Fc binding site is distant from the mechano-active site, the mechanism underlying the mechanical stabilization effect is likely a result of some sort of long-range coupling between the two sites. The enhancement in mechanical stability by the binding of Fc to GB1 is robust and can tolerate substantial changes to the protein structure. GB1 mutants NuG2 and Gc3b4 are two examples.^[90,91] Both NuG2 and Gc3b4 are computationally designed mutants of GB1. The first β -hairpin in NuG2 was redesigned as compared with wt GB1 and involves 11 mutations in total; in contrast, Gc3b4 involves seven point mutations in the hydrophobic core of GB1. Despite such a major change in their primary sequences, both NuG2 and Gc3b4 retain their binding capability to Fc. Similarly, the binding of Fc can significantly enhance their mechanical stability. For example, the mechanical unfolding force of NuG2 was doubled by the binding of Fc from ~105 to

~210 pN. For Gc3b4, the mechanical unfolding force was increased from 210 to ~300 pN. Using GB1 mutants with a different binding affinity to Fc, we also investigated whether the binding affinity is correlated with the amplitude of mechanical stability enhancement. We showed that, although the binding affinity of GB1 mutants to Fc varies by three orders of magnitude, the mechanical stability enhancement falls in a narrow range between 70 to 120 pN, which suggests that there is no direct correlation between mechanical stability enhancement and binding affinity. These results suggest that the binding affinity only affects the population of the GB1/hFc complex at a given concentration of hFc, but does not affect the intrinsic mechanical stability of the GB1/hFc complex.

The enhancement of mechanical stability by the binding of Fc ranges from 70 to 120 pN. Compared with the mechanical stability enhancement observed for Top7 by controlling the unfolding pathway by a disulfide bridge crosslink,^[40] the stabilization effect caused by a non-covalent interaction with Fc is truly substantial. Even more striking is that the Fc-binding site is quite distant from the mechano-active site and yet displays a significant mechanical stabilization. These results raise interesting questions as to what structural parameters determine the amplitude of the enhancement effect and how we can modulate this effect.

These studies demonstrate that ligand-binding and protein–protein interactions can serve as efficient and rational approaches to modulate the mechanical stability of a protein and will thus have tremendous potential in engineering smart proteins that are sensitive to and can be modulated by environmental stimuli. It is of note that such methodologies depend on specific protein–ligand systems and are not universal.^[66,67,89,92] In some cases, strong ligand binding or protein–protein interactions will not translate into mechanical stabilization. For example, a recent study on Im9 showed that the binding of a protein ligand E9, which has a dissociation constant of 10×10^{-15} M, does not change the mechanical stability of Im9.^[92] It is clear that many fundamental questions need to be addressed in order for us to use protein–protein interactions and ligand binding as general means in protein mechanics to tune the mechanical stability of proteins rationally and systematically. These questions include: What type of ligand binding can lead to mechanical stability enhancement? How does the long range coupling between the binding site and the mechano-active site work? What determines the amplitude for mechanical stability enhancement? Is it possible to develop more general and widely applicable ligand-binding schemes, such as metal chelation, to tune the mechanical stability of proteins?

4.6. Tuning the Mechanical Stability by Solvent Condition and Temperature

Solvent condition and temperature can also affect the thermodynamics of proteins based on different mechanisms,

and thus serve as convenient approaches to tune the mechanical stability of proteins. Chemical denaturants,^[56] osmolyte,^[93] and crowding agents (G. Yang, private communication) are representative means one can use to change the solvent condition and to affect the mechanical stability of proteins in solution.

We have systematically investigated the effect of denaturant on the mechanical stability of GB1.^[56] We found that increasing the concentration of chemical denaturant decreases the mechanical unfolding force of GB1 in a linear fashion. Detailed analysis of the mechanical unfolding rate constant revealed that the effect of a chemical denaturant on mechanical unfolding reduces the mechanical unfolding free energy barrier and thus increases the mechanical unfolding rate constant α . The relationship between α and denaturant concentration can be described by Equation (3):

$$RT \ln \alpha_0(\text{denaturant}) = RT \ln \alpha_0(\text{PBS}) + m_u[\text{GdmCl}] \quad (3)$$

which is similar to the relationship used to describe the effect of chemical denaturant on the chemical unfolding kinetics. The reduced mechanical stability is a result of a lowering of the mechanical unfolding barrier rather than a shift of the mechanical unfolding transition state. Because of this simple relationship between chemical denaturant concentration and mechanical unfolding force, it is possible to predict the mechanical unfolding force of GB1 at any given chemical denaturant concentration. However, since the chemical denaturant can only weaken a given protein both chemically and mechanically, the use of chemical denaturant to tune the mechanical stability of proteins is limited.

In contrast to the weakening effect of chemical denaturant, osmolyte molecules can stabilize the native conformations of proteins and enhance their thermodynamic stability. Recently, Fernandez and co-workers explored the use of glycerol to tune the mechanical stability of a small protein I27.^[93] They used single molecule AFM to examine the mechanical unfolding kinetics of I27 in various concentrations of glycerol. They observed that addition of glycerol to the solution increased the mechanical unfolding energy barrier in a glycerol-concentration dependent fashion. More interestingly, they also found that glycerol can participate in the formation of hydrogen bonds in the mechano-active site of I27 in the mechanical unfolding transition state, which leads to an increase of unfolding distance Δx_u from 0.25 nm for I27 in phosphate buffered saline (PBS) to 0.45 nm for I27 in buffer containing 30% or more glycerol. The complex role of glycerol in the mechanical unfolding gave rise to the surprising dependence of the mechanical unfolding force of I27 on the glycerol concentration: when the glycerol concentration is below 30%, the mechanical stability of I27 decreases from 200 pN (in PBS buffer) to ~150 pN (30% glycerol); when the glycerol concentration is higher than 30%, the mechanical stability of I27 increases as a function of the glycerol concentration to ~300 pN (100% glycerol). They interpret

this unusual dependence of mechanical stability on glycerol concentration using the relative change of unfolding energy barrier and unfolding distance Δx_u : at low glycerol concentration, the effect of increasing Δx_u overwhelms the effect of the increased mechanical unfolding energy barrier on the mechanical unfolding force; at a high concentration of glycerol, the increased mechanical unfolding barrier becomes the dominant factor in the mechanical unfolding of I27.

Similarly, the effect of temperature on the mechanical stability of proteins was also investigated. A thermo-softening effect was a general observation among different proteins being studied.^[94,95] The unfolding force of a given protein is higher at a lower temperature than that at a higher temperature.

5. Recombination Based Approach to Engineer Proteins of Novel Mechanical Properties: Towards Directed Evolution Approach for Protein Mechanics

Despite progress in the rational engineering of elastomeric proteins, our knowledge on the molecular determinants of the mechanical stability of proteins remains rather limited. For example, the giant muscle protein titin contains hundreds of Ig domains that possess highly homologous primary sequences and three-dimensional structures yet vastly different mechanical stability. Our current understanding still cannot comprehend the intricate molecular details that give rise to the diversity of mechanical stability of such highly homologous proteins. To circumvent these challenges, we have started to explore the use of recombination-based approaches to engineer elastomeric proteins with the hope to develop methods analogous to the directed evolution approaches widely used in the field of enzyme engineering.

Recombination is an important mechanism used by nature to use similar protein folds to accommodate distant sequences and different functions. Recombination offers the advantage of combining beneficial mutations from multiple parents into a single offspring and has been exploited extensively by nature during evolution to improve protein traits such as enzymatic activity. This idea has also been used extensively in laboratory based directed-evolution, and recombination has become one of the most important strategies in engineering proteins with novel functions.^[96,97] Using the Ig domains from the muscle protein titin, we started to explore the use of recombination of

protein fragments to explore the sequence space and engineer proteins of novel mechanical stability.^[74]

I27 and I32 are two well-characterized immunoglobulin domains^[33] from titin and are of distinct mechanical stability: I27 unfolds at ~ 200 pN while I32 unfolds at ~ 300 pN. I27 and I32 share high sequence homology and are ideal systems for protein recombination. Using these two proteins as parent templates, we have demonstrated the feasibility of recombining protein fragments to construct proteins of novel mechanical properties. We systematically interchanged the structural fragments between the two parent proteins and constructed a series of hybrid daughter proteins, four of which are shown in Figure 7. In these four hybrid proteins (I27-A'G-I32, I32-A'G-I27, I27-CDE-I32, and I32-CDE-I27), the force-bearing strands A'G, which is the mechano-active site, as well as the non-force bearing strands CDE were interchanged between I27 and I32, respectively. These four hybrid daughter proteins fold properly, as assessed by circular-dichroism spectroscopy, into well-defined β -sheet dominated structures that are similar to that of the

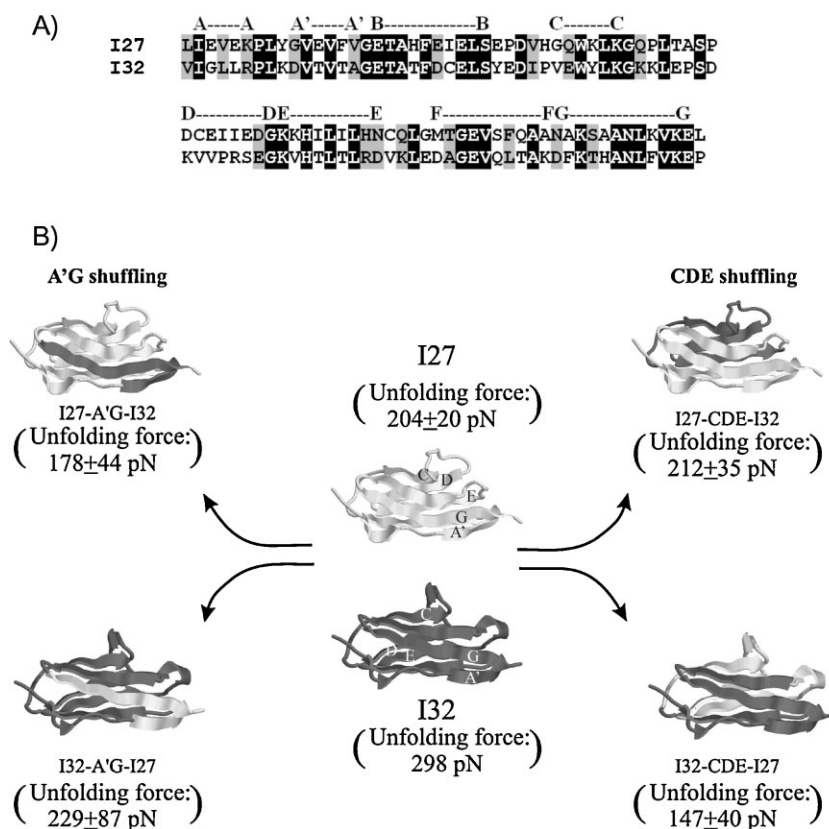


Figure 7. Engineering novel mechanical proteins by recombination of protein fragments from I27 and I32. A) Amino acid sequence alignment of I27 and I32. Grey shadings indicate homology, and inverse texts indicate identity B) Three dimensional structure and mechanical unfolding forces of hybrid Ig domains. Middle column shows the three dimensional structures of I27 (grey) and I32 (black). I32 structure was obtained by homology modeling. By interchanging the A' and G β strands between I27 and I32, hybrid proteins I27-A'G-I32 and I32-A'G-I27 were engineered (left column). Interchanging the C, D, and E β strands between I27 and I32 resulted in hybrid proteins I27-CDE-I32 and I32-CDE-I27 (right column). In the hybrid proteins, the fragments coming from the wild type I27 are shown in grey, while those from wild-type I32 are shown in black. The hybrid daughter proteins exhibit mechanical properties that are distinct from those of parent proteins. Adapted from^[74].

parent proteins. They also show mechanical stability that is distinct from that of both parent proteins (Fig. 7). Interestingly, two daughter proteins I27-CDE-I32 and I32-A'G-I27 exhibit mechanical stability that is higher than the parent protein I27, but lower than I32. None of the hybrid proteins show a mechanical stability that is higher than both parents. This study demonstrated the great potential of shuffling protein fragments among homologous parent proteins to engineer novel mechanically stable proteins. In principle, if one can build a library of hybrid proteins that is large enough, it is possible to obtain daughter proteins that are of diverse mechanical stability, including the ones that are mechanically stronger than both parent proteins. Such a library of proteins will make it possible to carry out statistical analysis to understand general rules about why some hybrid proteins are mechanically stable while others are not, despite their high sequence homology. Library construction and statistical analysis of this sort have already been successfully carried out in the field of enzyme engineering to engineer thermostable cytochrome P450s.^[98] Such pioneering work will certainly inspire similar experimental efforts in the field of protein mechanics.

Our recently expanded experimental efforts on I27 and I32 also revealed some complex issues in shuffling protein fragments to engineer proteins of novel mechanical properties. Recombining protein fragments from two homologous parent templates does not always generate properly folded daughter proteins, despite their high sequence homology. It is clear that random recombination may disrupt interactions that are present in the parent proteins and are critical for the folding of the protein. Disruption of such interactions can lead to misfolding of the constructed hybrid daughter proteins. Developing efficient ways to screen hybrid proteins that are properly folded will be an important task towards generating a large library of hybrid proteins. Alternatively, a computational design guided recombination approach will have to be employed, just as those developed by Arnold and co-workers to guide the recombination efforts.^[97]

Moreover, to develop the recombination approach into a directed-evolution technique, an even greater challenge is present: developing efficient screening assays to select hybrid proteins with desirable mechanical stability. Currently, determining the mechanical stability of the hybrid daughter proteins still relies on the use of single molecule AFM and construction of polyproteins, which are not realistic when handling a large library of proteins. Developing high throughput and alternative approaches to determine the mechanical stability of proteins will be an important task for future endeavors.

6. Future Perspective

6.1. Promises and Challenges

Since the first single molecule force spectroscopy experiments on titin carried out a decade ago, single protein mechanics and engineering has made tremendous progress and

has evolved into a full fledged field of inquiry. Using single molecule force spectroscopy techniques, the mechanical properties of individual proteins can now be examined in great detail; using protein engineering techniques, proteins with novel design to entail tailored mechanical properties can be engineered and examined at the single molecule level. This is a burgeoning field full of promises and challenges, and greater discoveries in this field are yet to be made.

The powerful combination of single molecule AFM with protein engineering techniques has brought us a step closer towards understanding the molecular determinants of mechanical stability of proteins, and a step closer towards tuning the mechanical properties of proteins in a rational fashion. The inclusion of non-mechanical proteins in the toolbox of elastomeric proteins has significantly expanded the range of proteins that can be utilized in mechanical application, and provides the foundation for constructing multifunctional nanomechanical materials and devices. For example, the investigation of the mechanical properties of green fluorescent protein (GFP) and the construction of GFPs with well defined mechanical properties may provide the possibility of using GFP as in-situ force sensors to directly report the stretching force in biomaterials as well as in living systems. Developing smart elastomeric materials should also be within the reach in the near future. The responsiveness of the mechanical stability of proteins to ligand binding, protein-protein interactions, and light^[41] have opened the possibility to develop smart elastomeric materials that are sensitive to and can be regulated by external stimuli.

The mechanical engineering of elastomeric proteins has become a reality for protein engineers. However, the tools and methodologies available to protein engineers are still rather limited. One of the main obstacles that remains is the lack of a full understanding of the molecular determinants of mechanical stability for proteins. General methodologies that allow for the rational tuning of the mechanical stability of a wide range of proteins are yet to be developed. In contrast, decades of work in enzyme engineering have unveiled many successful strategies to enhance the thermodynamic stability of proteins. Owing to the difference in thermodynamic stability and mechanical stability, these techniques cannot be directly put to use in protein mechanics. However, their underlying fundamental thermodynamic principles are universal and should offer similar tricks one can employ in protein mechanics. The challenge is how to adapt these successful strategies for the use in protein mechanics, e.g., how to apply these strategies to the mechanical unfolding pathways of proteins in order to preferentially stabilize the native state over the mechanical unfolding transition state.

Another challenge in protein mechanics is the use of computational methodology to facilitate the design of elastomeric proteins with tailored mechanical properties. The computational design of proteins with atomic level accuracy has achieved tremendous progress and success in recent years. A recent example has demonstrated the *ab initio* design of fully functional enzymes that are able to use

substrates that natural enzymes cannot.^[99] Considering that the mechanical stability of proteins is a property dictated by the overall three-dimensional structure, designing novel proteins with well-defined mechanical stability should be within the reach of the current computational biology.

6.2. From Single Molecules to Protein-Based Materials

The developments in single protein mechanics not only provide new insights into the mechanical design of proteins at the single molecule level, but also open up a completely new perspective for material sciences. The mechanical properties of biomaterials/machinery, such as spider silk and muscle fibers, are generally understood at the macroscopic level, and the mechanical properties of the constituting elastomeric proteins are thus deduced from the available ensemble of experimental data. The ability to characterize the mechanical properties of elastomeric proteins at the single molecule level and design novel elastomeric proteins with well-defined mechanical properties will enable the construction of a new generation of protein-based biomaterials by a bottom-up approach using well-defined building blocks. Such materials will provide the possibility to establish the relationship between the macroscopic mechanical performance and the mechanical features at the single molecule level. Similar ideals have recently been employed in polymer chemistry to synthesize new polymers with built-in sacrificial bonds to increase the mechanical strength of the synthetic polymer materials.^[100] For proteins, the challenge will be to develop methodologies to effectively integrate nanometer-scaled building blocks of elastomeric proteins into macroscopic materials with controlled organization of constituting elastomeric proteins. As the first attempt to build macroscopic materials with elastomeric proteins, recently we have started to explore the use of our engineered artificial polyprotein (GB1)₈ to construct novel protein-based hydrogels.^[101] Since the mechanical properties of building block proteins can be tailored at the single molecule level, exciting opportunities are present to test the relationship between single molecule mechanics and macroscopic mechanical performance. We anticipate that such artificial polyprotein-based biomaterials will lead to new materials of well-defined and novel mechanical properties that will find applications in fields that range from material sciences to biomedical engineering.^[102,103] 'Mechanical engineering' of proteins at the single molecule level has led us into a new era of biomaterial research, and surely new exciting discoveries and surprises are yet to come!

Received: April 8, 2008

Published online: September 1, 2008

- [1] G. Bao, S. Suresh, *Nat. Mater.* **2003**, 2, 715.
- [2] J. Gosline, M. Lillie, E. Carrington, P. Guerette, C. Ortlepp, K. Savage, *Phil. Trans. R. Soc. London Ser. B: Biol. Sci.* **2002**, 357, 121.
- [3] A. S. Tatham, P. R. Shewry, *Trends BioChem. Sci.* **2000**, 25, 567.

- [4] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, *Science* **1997**, 276, 1109.
- [5] S. Labeit, B. Kolmerer, *Science* **1995**, 270, 293.
- [6] B. L. Smith, T. E. Schaffer, M. Viani, J. B. Thompson, N. A. Frederick, J. Kindt, A. Belcher, G. D. Stucky, D. E. Morse, P. K. Hansma, *Nature* **1999**, 399, 761.
- [7] A. F. Oberhauser, P. E. Marszalek, H. P. Erickson, J. M. Fernandez, *Nature* **1998**, 393, 181.
- [8] T. Ohashi, D. P. Kiehart, H. P. Erickson, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 2153.
- [9] P. A. Guerette, D. G. Ginzinger, B. H. Weber, J. M. Gosline, *Science* **1996**, 272, 112.
- [10] N. Becker, E. Oroudjev, S. Mutz, J. P. Cleveland, P. K. Hansma, C. Y. Hayashi, D. E. Makarov, H. G. Hansma, *Nat. Mater.* **2003**, 2, 278.
- [11] C. M. Elvin, A. G. Carr, M. G. Huson, J. M. Maxwell, R. D. Pearson, T. Vuocolo, N. E. Liyou, D. C. Wong, D. J. Merritt, N. E. Dixon, *Nature* **2005**, 437, 999.
- [12] M. S. Kellermayer, S. B. Smith, H. L. Granzier, C. Bustamante, *Science* **1997**, 276, 1112.
- [13] L. Tskhovrebova, J. Trinick, J. A. Sleep, R. M. Simmons, *Nature* **1997**, 387, 308.
- [14] W. A. Linke, A. Grutzner, *Pflugers Arch.* **2008**, 456, 101.
- [15] A. F. Oberhauser, M. Carrion-Vazquez, *J. Biol. Chem.* **2008**, 283, 6617.
- [16] H. Li, *Org. Biomol. Chem.* **2007**, 5, 3399.
- [17] D. W. Urry, *Trends Biotechnol.* **1999**, 17, 249.
- [18] M. Carrion-Vazquez, A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. Li, J. M. Fernandez, *Prog. Biophys. Mol. Biol.* **2000**, 74, 63.
- [19] D. J. Brockwell, *Curr. Nanosci.* **2007**, 3, 3.
- [20] T. E. Fisher, A. F. Oberhauser, M. Carrion-Vazquez, P. E. Marszalek, J. M. Fernandez, *Trends Biochem. Sci.* **1999**, 24, 379.
- [21] X. W. Zhuang, M. Rief, *Curr. Opin. Struct. Biol.* **2003**, 13, 88.
- [22] M. Carrion-Vazquez, A. Oberhauser, H. Diez, R. Hervas, J. Oroz, J. Fernandez, A. Martinez-Martin, in *Advanced Techniques in Biophysics*, (Eds: J. Arrondo, A. Alonso), Springer, Berlin **2006**, p. 163.
- [23] G. Lee, K. Abdi, Y. Jiang, P. Michaely, V. Bennett, P. E. Marszalek, *Nature* **2006**, 440, 246.
- [24] M. Sotomayor, K. Schulten, *Science* **2007**, 316, 1144.
- [25] E. Paci, M. Karplus, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 6521.
- [26] R. Rounsevell, J. R. Forman, J. Clarke, *Methods* **2004**, 34, 100.
- [27] L. Li, S. Wetzel, A. Pluckthun, J. M. Fernandez, *Biophys. J.* **2006**, 90, L30.
- [28] P. Carl, C. H. Kwok, G. Manderson, D. W. Speicher, D. E. Discher, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 1565.
- [29] B. Bullard, T. Garcia, V. Benes, M. C. Leake, W. A. Linke, A. F. Oberhauser, *Proc. Natl. Acad. Sci. USA* **2006**, 103, 4451.
- [30] C. Cecconi, E. A. Shank, C. Bustamante, S. Marqusee, *Science* **2005**, 309, 2057.
- [31] C. L. Chyan, F. C. Lin, H. Peng, J. M. Yuan, C. H. Chang, S. H. Lin, G. Yang, *Biophys. J.* **2004**, 87, 3995.
- [32] I. Schwaiger, C. Sattler, D. R. Hostetter, M. Rief, *Nat. Mater.* **2002**, 1, 232.
- [33] H. Li, W. A. Linke, A. F. Oberhauser, M. Carrion-Vazquez, J. G. Kerkvliet, H. Lu, P. E. Marszalek, J. M. Fernandez, *Nature* **2002**, 418, 998.
- [34] H. Li, J. M. Fernandez, *J. Mol. Biol.* **2003**, 334, 75.
- [35] H. Li, A. F. Oberhauser, S. B. Fowler, J. Clarke, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 6527.
- [36] H. Li, A. F. Oberhauser, S. D. Redick, M. Carrion-Vazquez, H. P. Erickson, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 10682.
- [37] G. Yang, C. Cecconi, W. A. Baase, I. R. Vetter, W. A. Breyer, J. A. Haack, B. W. Matthews, F. W. Dahlquist, C. Bustamante, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 139.

- [38] D. J. Brockwell, G. S. Beddard, E. Paci, D. K. West, P. D. Olmsted, D. A. Smith, S. E. Radford, *Biophys. J.* **2005**, *89*, 506.
- [39] H. Dietz, M. Rief, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16 92.
- [40] D. Sharma, O. Perisic, Q. Peng, Y. Cao, C. Lam, H. Lu, H. Li, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9278.
- [41] J. M. Zhao, H. Lee, R. A. Nome, S. Majid, N. F. Scherer, W. D. Hoff, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11 561.
- [42] Y. Cao, C. Lam, M. Wang, H. Li, *Angew. Chem. Int. Ed.* **2006**, *45*, 642.
- [43] Y. Cao, H. Li, *Nat. Mater.* **2007**, *6*, 109.
- [44] M. Carrion-Vazquez, A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3694.
- [45] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12 724.
- [46] M. Rief, F. Oesterhelt, B. Heymann, H. E. Gaub, *Science* **1997**, *275*, 1295.
- [47] A. F. Oberhauser, P. K. Hansma, M. Carrion-Vazquez, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 468.
- [48] M. Schlierf, H. Li, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7299.
- [49] M. J. Wang, Y. Cao, H. B. Li, *Polymer* **2006**, *47*, 2548.
- [50] G. I. Bell, *Science* **1978**, *200*, 618.
- [51] E. Evans, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 105.
- [52] E. Evans, K. Ritchie, *Biophys. J.* **1997**, *72*, 1541.
- [53] H. Li, M. Carrion-Vazquez, A. F. Oberhauser, P. E. Marszalek, J. M. Fernandez, *Nat. Struct. Biol.* **2000**, *7*, 1117.
- [54] D. J. Brockwell, G. S. Beddard, J. Clarkson, R. C. Zinober, A. W. Blake, J. Trinick, P. D. Olmsted, D. A. Smith, S. E. Radford, *Biophys. J.* **2002**, *83*, 458.
- [55] P. M. Williams, S. B. Fowler, R. B. Best, J. L. Toca-Herrera, K. A. Scott, A. Steward, J. Clarke, *Nature* **2003**, *422*, 446.
- [56] Y. Cao, H. Li, *J. Mol. Biol.* **2008**, *375*, 316.
- [57] P. E. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, J. M. Fernandez, *Nature* **1999**, *402*, 100.
- [58] K. A. Scott, A. Steward, S. B. Fowler, J. Clarke, *J. Mol. Biol.* **2002**, *315*, 819.
- [59] M. Carrion-Vazquez, H. Li, H. Lu, P. E. Marszalek, A. F. Oberhauser, J. M. Fernandez, *Nat. Struct. Biol.* **2003**, *10*, 738.
- [60] D. J. Brockwell, E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham, S. E. Radford, *Nat. Struct. Biol.* **2003**, *10*, 731.
- [61] D. K. West, D. J. Brockwell, P. D. Olmsted, S. E. Radford, E. Paci, *Biophys. J.* **2006**, *90*, 287.
- [62] J. R. Forman, J. Clarke, *Curr. Opin. Struct. Biol.* **2007**, *17*, 58.
- [63] D. S. Goodsell, *Bionanotechnology*, Wiley-Liss, Hoboken **2004**.
- [64] A. Ikai, R. Afrin, H. Sekiguchi, *Curr. Nanosci.* **2007**, *3*, 17.
- [65] R. B. Best, B. Li, A. Steward, V. Daggett, J. Clarke, *Biophys. J.* **2001**, *81*, 2344.
- [66] J. P. Junker, K. Hell, M. Schlierf, W. Neupert, M. Rief, *Biophys. J.* **2005**, *89*, L46.
- [67] A. J. Wilcox, J. Choy, C. Bustamante, A. Matouschek, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15435.
- [68] Q. Peng, H. Li, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1885.
- [69] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, *Biophys. J.* **1998**, *75*, 662.
- [70] H. Lu, K. Schulten, *Biophys. J.* **2000**, *79*, 51.
- [71] D. K. Klimov, D. Thirumalai, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7254.
- [72] J. I. Sulkowska, M. Cieplak, *Biophys. J.* **2008**, *94*, 6.
- [73] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, *Science* **2003**, *302*, 1364.
- [74] D. Sharma, Y. Cao, H. Li, *Angew. Chem. Int. Ed.* **2006**, *45*, 5633.
- [75] R. Perez-Jimenez, S. Garcia-Manyes, S. R. Ainarapu, J. M. Fernandez, *J. Biol. Chem.* **2006**, *281*, 40 010.
- [76] R. B. Best, S. B. Fowler, J. L. Herrera, A. Steward, E. Paci, J. Clarke, *J. Mol. Biol.* **2003**, *330*, 867.
- [77] S. B. Fowler, R. B. Best, J. L. Toca Herrera, T. J. Rutherford, A. Steward, E. Paci, M. Karplus, J. Clarke, *J. Mol. Biol.* **2002**, *322*, 841.
- [78] L. Li, H. H. Huang, C. L. Badilla, J. M. Fernandez, *J. Mol. Biol.* **2005**, *345*, 817.
- [79] I. Schwaiger, A. Kardinal, M. Schleicher, A. A. Noegel, M. Rief, *Nat. Struct. Mol. Biol.* **2004**, *11*, 81.
- [80] S. P. Ng, K. S. Billings, T. Ohashi, M. D. Allen, R. B. Best, L. G. Randles, H. P. Erickson, J. Clarke, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9633.
- [81] Y. Cao, T. Yoo, S. Zhuang, H. Li, *J. Mol. Biol.* **2008**, *378*, 1132.
- [82] E. Paci, M. Karplus, *J. Mol. Biol.* **1999**, *288*, 441.
- [83] M. Gao, D. Craig, V. Vogel, K. Schulten, *J. Mol. Biol.* **2002**, *323*, 939.
- [84] H. Li, H.-C. Wang, Y. Cao, D. Sharma, M. Wang, *J. Mol. Biol.* **2008**, *379*, 871.
- [85] S. R. Ainarapu, J. Brujic, H. H. Huang, A. P. Wiita, H. Lu, L. Li, K. A. Walther, M. Carrion-Vazquez, H. Li, J. M. Fernandez, *Biophys. J.* **2007**, *92*, 225.
- [86] S. M. Malakauskas, S. L. Mayo, *Nat. Struct. Biol.* **1998**, *5*, 470.
- [87] J. M. Sanchez-Ruiz, *Biophys. Chem.* **2007**, *126*, 43.
- [88] H. Fukada, J. M. Sturtevant, F. A. Quiocho, *J. Biol. Chem.* **1983**, *258*, 13 193.
- [89] S. R. Ainarapu, L. Li, C. L. Badilla, J. M. Fernandez, *Biophys. J.* **2005**, *89*, 3337.
- [90] Y. Cao, M. M. Balamurali, D. Sharma, H. Li, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15 677.
- [91] Y. Cao, T. Yoo, S. Zhuang, H. Li, *J. Mol. Biol.* **2008**, *378*, 1132.
- [92] E. Hann, N. Kirkpatrick, C. Kleanthous, D. A. Smith, S. E. Radford, D. J. Brockwell, *Biophys. J.* **2007**, *92*, L79.
- [93] L. Dougan, G. Feng, H. Lu, J. M. Fernandez, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3185.
- [94] R. Law, G. Liao, S. Harper, G. Yang, D. W. Speicher, D. E. Discher, *Biophys. J.* **2003**, *85*, 3286.
- [95] M. Schlierf, M. Rief, *J. Mol. Biol.* **2005**, *354*, 497.
- [96] J. D. Bloom, M. M. Meyer, P. Meinhold, C. R. Otey, D. MacMillan, F. H. Arnold, *Curr. Opin. Struct. Biol.* **2005**, *15*, 447.
- [97] C. A. Voigt, C. Martinez, Z. G. Wang, S. L. Mayo, F. H. Arnold, *Nat. Struct. Biol.* **2002**, *9*, 553.
- [98] Y. Li, D. A. Drummond, A. M. Sawayama, C. D. Snow, J. D. Bloom, F. H. Arnold, *Nat. Biotechnol.* **2007**, *25*, 1051.
- [99] L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas, 3rd, D. Hilvert, K. N. Houk, B. L. Stoddard, D. Baker, *Science* **2008**, *319*, 1387.
- [100] Z. Guan, J. T. Roland, J. Z. Bai, S. X. Ma, T. M. McIntire, M. Nguyen, *J. Am. Chem. Soc.* **2004**, *126*, 2058.
- [101] Y. Cao, H. Li, *Chem. Commun.* **2008**, In press.
- [102] R. Langer, D. A. Tirrell, *Nature* **2004**, *428*, 487.
- [103] X. Y. Wang, H. J. Kim, C. Wong, C. Vepari, A. Matsumoto, D. L. Kaplan, *Mater. Today* **2006**, *9*, 44.